# HL7 Messaging and caBIG

- Heath Level Seven (**HL7**) is an ANSI- accredited Standards Developing Organization (SDO) operating in healthcare arena

- HL7 serves as a way for inherently disparate applications and data architectures operating in a heterogeneous system environment to communicate with each other.

- Designed to support distributed environment where data resides in departmental systems

- Recommended by the NCVHS and CHI as the messaging standard for electronic exchange of clinical data.

- Over 90% of US hospitals have implemented some version of **HL7** messages (almost all 2.x)

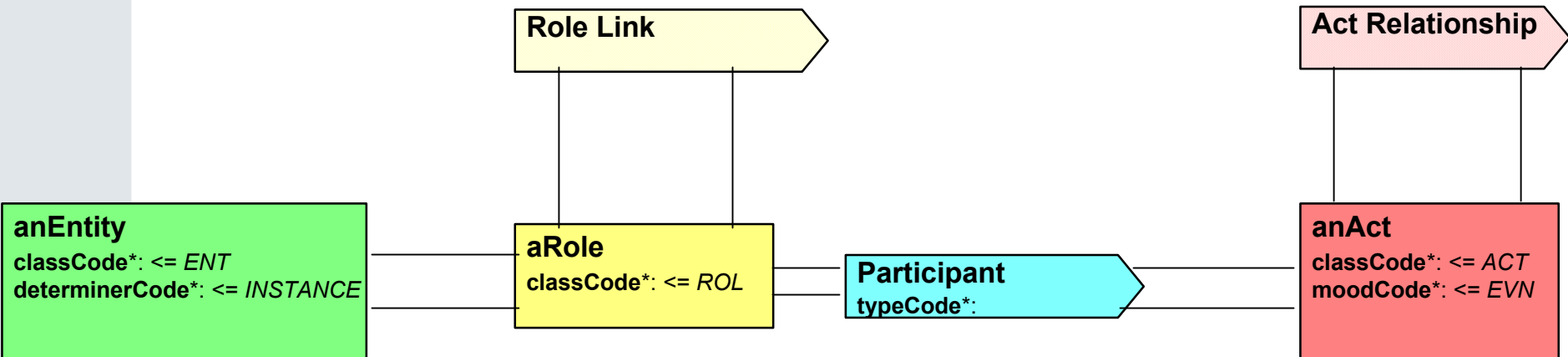- Significant body of work that can be leveraged for caBIG uses
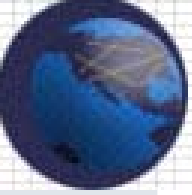
- Internationalization
- OO and UML RIM based
- Standardization of vocabulary domains
- Reduced optionality for improved semantic consistency
- V2.x Functional Compatibility

- **Information Model** - Describes the types of information models in HL7's Version 3 standard, including the Reference Information Model (RIM).
- **Vocabulary** - Discusses the use of controlled vocabulary within the HL7 V3 Specifications and the HL7 Vocabulary Technical Committee's principles and approach to vocabulary use and management.
- **Implementation Technology Specification (ITS)** - Describes the Implementation Technology Specification document and its function.
- **Data Types** - Describes the underlying principles of the V3 data types.
- **Common Message Element Types (CMETs)** - Describes the purpose, definition, and use of CMETs.

Requirement	V3 Guide Construct(s)	Level
A means of providing context to the definitions of trigger events	Storyboards, state diagrams	Informative
A means of specifying the information content of messages through a common information model that clarifies the definitions and ensures that they are used consistently across all V3 messages defined by all Technical Committees	Reference Information Model	Reference
A means of specifying responsibilities of the senders and receivers of messages	Interaction model	Normative
A common description of the exact fields of a message and their grouping, sequence, optionality, and cardinality	Hierarchical Message Descriptions	Normative
Separate syntax specifications, describing the algorithms used to encode and transmit the messages in an XML based character stream syntax.	Implementation Technology Specifications	Normative

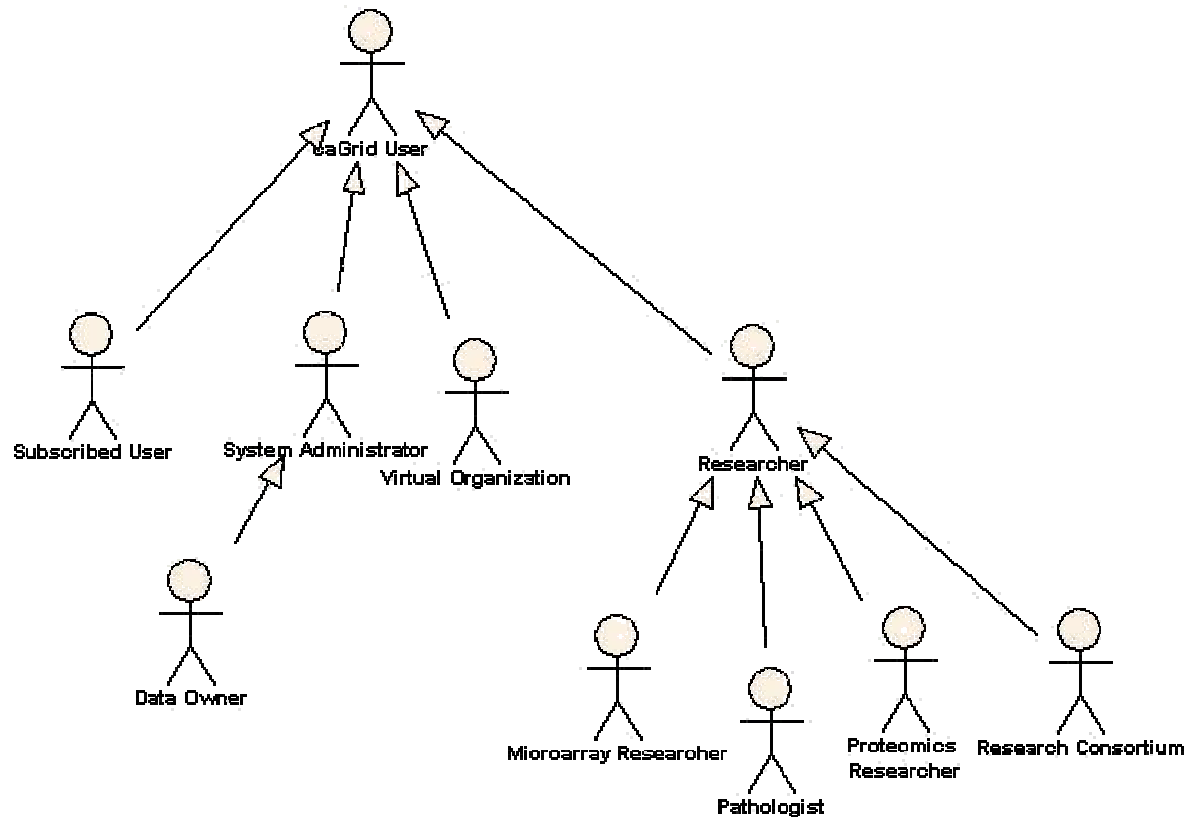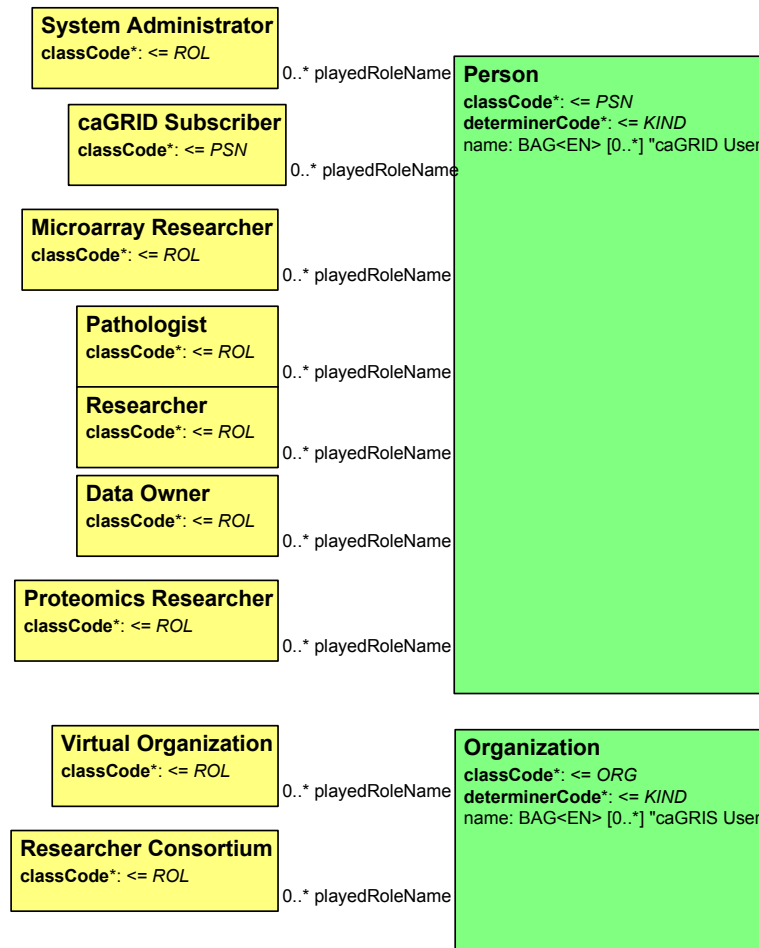caBIG — cancer Biomedical Informatics Grid

Role Link

Act Relationship

anEntity
classCode*: <= ENT
determinerCode*: <= INSTANCE

aRole
classCode*: <= ROL

Participant
typeCode*:

anAct
classCode*: <= ACT
moodCode*: <= EVN

cancer Biomedical Informatics Grid

caBIG

System Administrator
classCode*: <= *ROL*

0..* playedRoleName

caGRID Subscriber
classCode*: <= *PSN*

0..* playedRoleName

Microarray Researcher
classCode*: <= *ROL*

0..* playedRoleName

Pathologist
classCode*: <= *ROL*

0..* playedRoleName

Researcher
classCode*: <= *ROL*

0..* playedRoleName

Data Owner
classCode*: <= *ROL*

0..* playedRoleName

Proteomics Researcher
classCode*: <= *ROL*

0..* playedRoleName

Person
classCode*: <= *PSN*
determinerCode*: <= *KIND*
name: BAG<EN> [0..*] "caGRID User"

Virtual Organization
classCode*: <= *ROL*

0..* playedRoleName

Researcher Consortium
classCode*: <= *ROL*

0..* playedRoleName

Organization
classCode*: <= *ORG*
determinerCode*: <= *KIND*
name: BAG<EN> [0..*] "caGRIS User"

- Root Element - The root element of the XML instance contains information necessary for its proper validation.

- Control Act - The message control act is a wrapper around the actual message. This indicates what response to the message is expected.

- Message Body - payload

- Clinical Trials area actively developing messages for:
  - Annotated ECG
  - Drug Product Stability Report
  - Adverse Event Reporting
  - Clinical Trial Lab Data

- Data types and Vocabulary components
- **HL7** clinical domains including:
  - Regulated Studies (Clinical Trials)
  - Public Health (Adverse Event)
  - Patient Administration
  - Infrastructure and implementation artifacts

DataValue	ANY	Defines the basic properties of every data value. This is an abstract type, meaning that no value can be just a data value without belonging to any concrete type. Every concrete type is a specialization of this general abstract DataValue type.
Boolean	BL	The Boolean type stands for the values of two-valued logic. A Boolean value can be either *true* or *false*, or, as any other value may be NULL.
BooleanNonNull	BN	The BooleanNonNull constrains the boolean type so that the value may not be NULL. This type is created for use within the data types specification where it is not appropriate for a null value to be used
Encapsulated Data	ED	Data that is primarily intended for human interpretation or for further machine processing outside the scope of HL7. This includes unformatted or formatted written language, multimedia data, or structured information in as defined by a different standard (e.g., XML-signatures.) Instead of the data itself, an ED may contain only a reference (see TEL.) Note that the ST data type is a specialization of the ED data type when the ED media type is text/plain.
Character String	ST	The character string data type stands for text data, primarily intended for machine processing (e.g., sorting, querying, indexing, etc.) Used for names, symbols, and formal expressions.

Concept Descriptor	CD	A concept descriptor represents any kind of concept usually by giving a code defined in a code system. A concept descriptor can contain the original text or phrase that served as the basis of the coding and one or more translations into different coding systems. A concept descriptor can also contain qualifiers to describe, e.g., the concept of a "left foot" as a postcoordinated term built from the primary code "FOOT" and the qualifier "LEFT". In cases of an exceptional value, the concept descriptor need not contain a code but only the original text describing that concept.
Coded Simple Value	CS	Coded data in its simplest form, where only the code is not predetermined. The code system and code system version are fixed by the context in which the CS value occurs. CS is used for coded attributes that have a single HL7-defined value set.
Coded Ordinal	CO	Coded data, where the domain from which the codeset comes is ordered. The Coded Ordinal data type adds semantics related to ordering so that models that make use of such domains may introduce model elements that involve statements about the order of the terms in a domain.
Coded With Equivalents	CE	Coded data that consists of a coded value (CV) and, optionally, coded value(s) from other coding systems that identify the same concept. Used when alternative codes may exist.
Character String with Code	SC	A character string that optionally may have a code attached. The text must always be present if a code is present. The code is often a local code.
Instance Identifier	II	An identifier that uniquely identifies a thing or object. Examples are object identifier for HL7 RIM objects, medical record number, order id, service catalog item id, Vehicle Identification Number (VIN), etc. Instance identifiers are defined based on ISO object identifiers.

Telecommunication Address	TEL	A telephone number (voice or fax), e-mail address, or other locator for a resource mediated by telecommunication equipment. The address is specified as a Universal Resource Locator (URL) qualified by time specification and use codes that help in deciding which address to use for a given time and purpose.
Postal Address	AD	Mailing and home or office addresses. A sequence of address parts, such as street or post office Box, city, postal code, country, etc.
Entity Name	EN	A name for a person, organization, place or thing. A sequence of name parts, such as given name or family name, prefix, suffix, etc. Examples for entity name values are "Jim Bob Walton, Jr.", "Health Level Seven, Inc.", "Lake Tahoe", etc. An entity name may be as simple as a character string or may consist of several entity name parts, such as, "Jim", "Bob", "Walton", and "Jr.", "Health Level Seven" and "Inc.", "Lake" and "Tahoe".
Trivial Name	TN	A restriction of entity name that is effectively a simple string used for a simple name for things and places.
Person Name	PN	An Entity Name used when the named Entity is a Person. A sequence of name parts, such as given name or family name, prefix, suffix, etc. A name part is a restriction of entity name part that only allows those entity name parts qualifiers applicable to person names. Since the structure of entity name is mostly determined by the requirements of person name, the restriction is very minor.
Organization Name	ON	An Entity Name used when the named Entity is an Organization. A sequence of name parts.

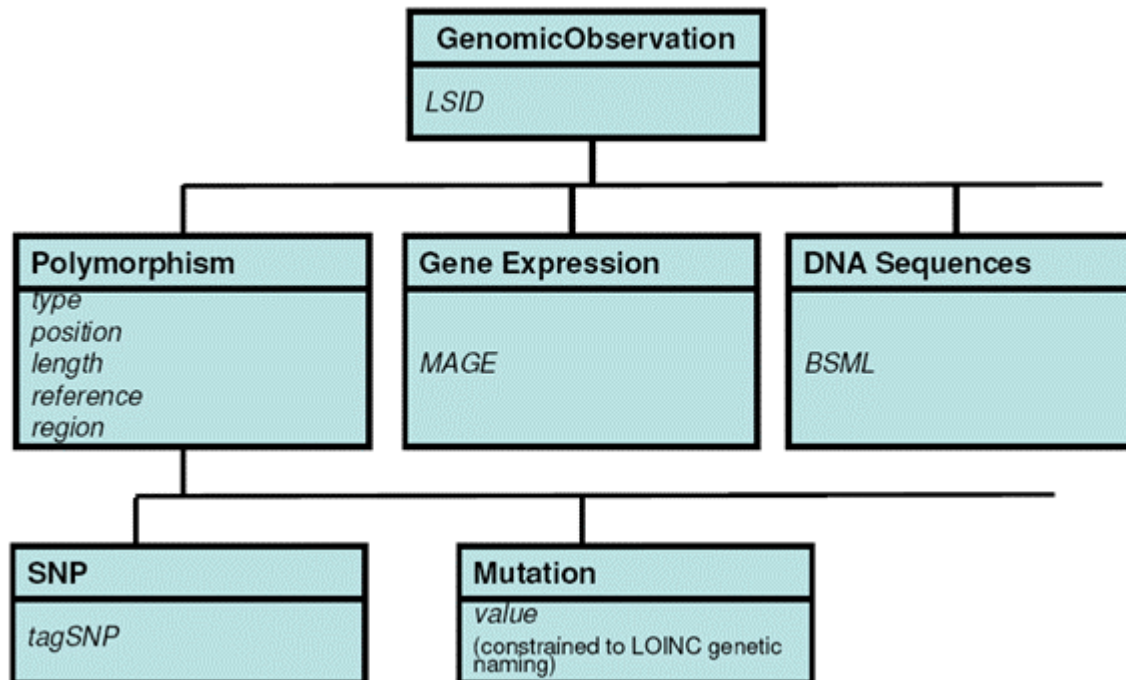| Integer Number	INT	Integer numbers (-1,0,1,2, 100, 3398129, etc.) are precise numbers that are results of counting and enumerating. Integer numbers are discrete, the set of integers is infinite but countable. No arbitrary limit is imposed on the range of integer numbers. Two NULL flavors are defined for the positive and negative infinity.
Real Number	REAL	Fractional numbers. Typically used whenever quantities are measured, estimated, or computed from other real numbers. The typical representation is decimal, where the number of significant decimal digits is known as the precision.
Ratio	RTO	A quantity constructed as the quotient of a numerator quantity divided by a denominator quantity. Common factors in the numerator and denominator are not automatically cancelled out. The *RTO* data type supports titers (e.g., "1:128") and other quantities produced by laboratories that truly represent ratios. Ratios are not simply "structured numerics", particularly blood pressure measurements (e.g. "120/60") are not ratios. In many cases the REAL should be used instead of the *RTO*.
Physical Quantity	PQ	A dimensioned quantity expressing the result of measuring.
Monetary Amount	MO	A monetary amount is a quantity expressing the amount of money in some currency. Currencies are the units in which monetary amounts are denominated in different economic regions. While the monetary amount is a single kind of quantity (money) the exchange rates between the different units are variable. This is the principle difference between physical quantity and monetary amounts, and the reason why currency units are not physical units.
Point in Time	TS	A quantity specifying a point on the axis of natural time. A point in time is most often represented as a calendar expression.

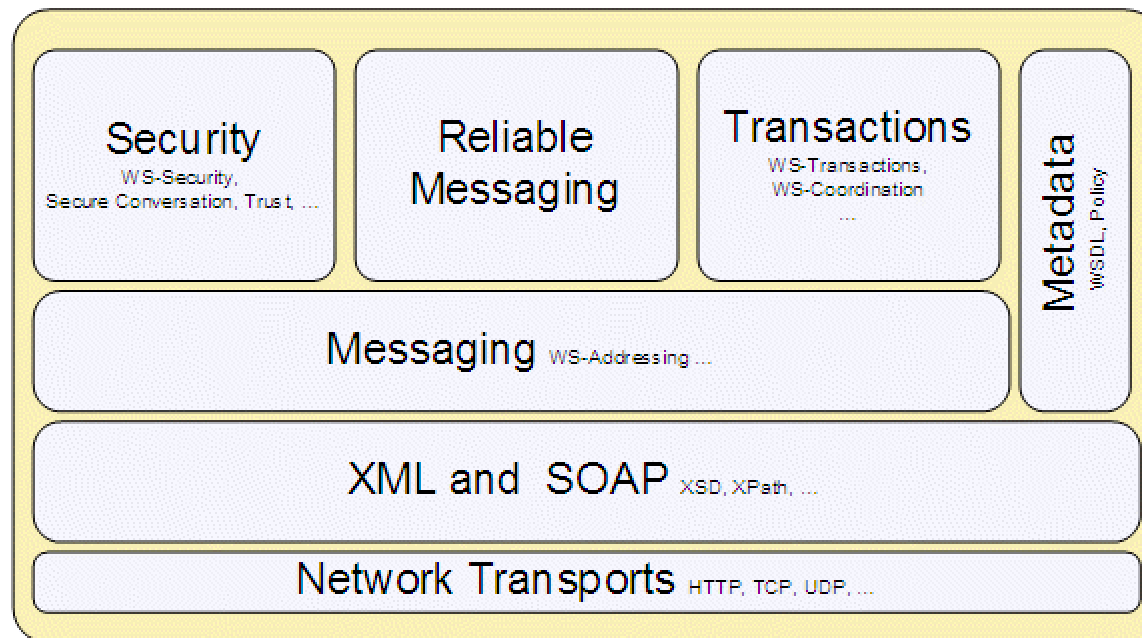| Set	SET	A value that contains other distinct values in no particular order.
Sequence	LIST	A value that contains other discrete values in a defined sequence.
Bag	BAG	An unordered collection of values, where each value can be contained more than once in the collection.
Interval	IVL	A set of consecutive values of an ordered base data type.
History	HIST	A set of data values that have a a valid-time property and thus conform to the history item (HXIT) type. The history information is not limited to the past; expected future values can also appear.
Uncertain Value - Probabilistic	UVP	A generic data type extension used to specify a probability expressing the information producer's belief that the given value holds.
Periodic Interval of Time	PIVL	An interval of time that recurs periodically. Periodic intervals have two properties, phase and period. The phase specifies the "interval prototype" that is repeated every period.
Event-Related Periodic Interval of Time	EIVL	Specifies a periodic interval of time where the recurrence is based on activities of daily living or other important events that are time-related but not fully determined by time.
General Timing Specification	GTS	A set of points in time, specifying the timing of events and actions and the cyclical validity-patterns that may exist for certain kinds of information, such as phone numbers (evening, daytime), addresses (so called "snowbirds," residing closer to the equator during winter and farther from the equator during summer) and office hours.
Parametric Probability Distribution	PPD	A generic data type extension specifying uncertainty of quantitative data using a distribution function and its parameters. Aside from the specific parameters of the distribution, a mean (expected value) and standard deviation is always given to help maintain a minimum layer of interoperability if receiving applications cannot deal with a certain probability distribution.

```xml
<xsd:complexTypename="CD">
  <xsd:complexContent>
    <xsd:extensionbase="ANY">
      <xsd:sequence>
        <xsd:elementname="originalText" type="ED" ... />
        <xsd:elementname="qualifier" type="CR" ... />
        <xsd:elementname="translation" type="CD" ... />
      </xsd:sequence>
      <xsd:attributename="code" type="cs" ... />
      <xsd:attributename="codeSystem" type="uid" ... />
      <xsd:attributename="codeSystemName" type="st" ... />
      <xsd:attributename="codeSystemVersion" type="st" ... />
      <xsd:attributename="displayName" type="st" ... />
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

caBIG
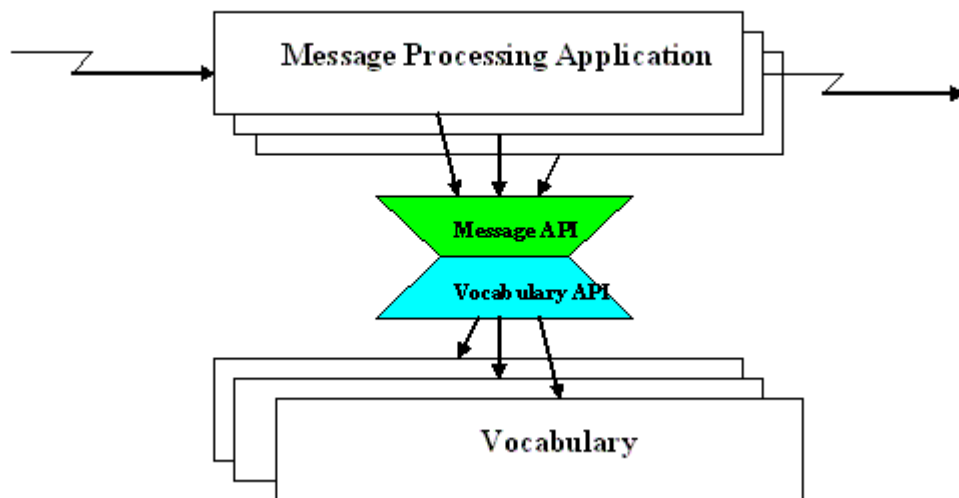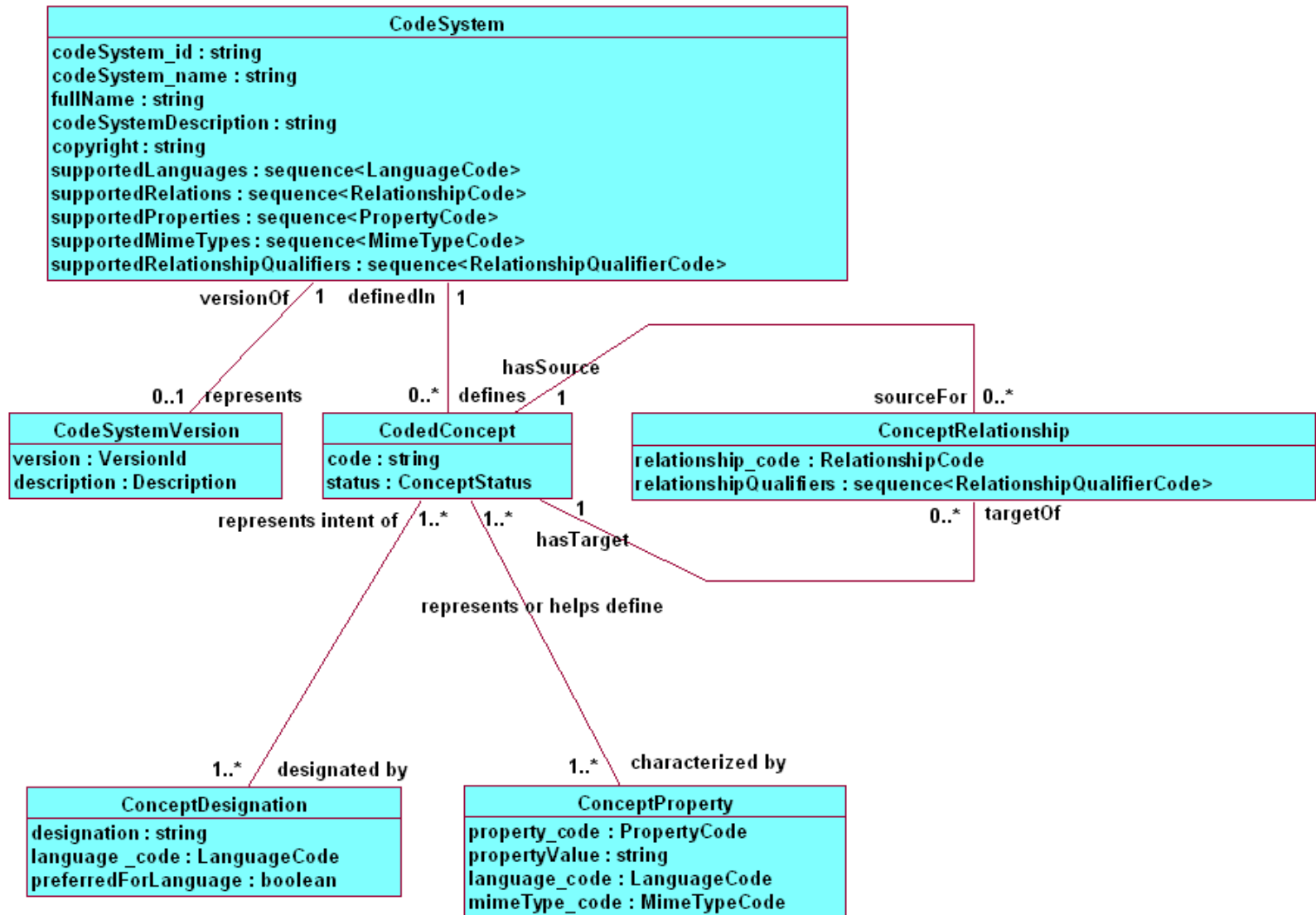
cancer Biomedical
Informatics Grid

SOPA 1.1 , WSDL 1.1

Defines an Application Programming Interface (API) that can be used by HL7 Version 3 software when accessing terminological content

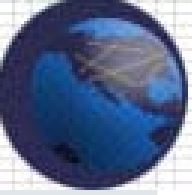Takes a resource identifier and concept code as input and returns a true/false value

- XML Implementation Technology Specification for V3 Structures

- UML Implementation Technology Specification - Data Types

- Clinical Document Architecture – May be useful in limited scope (Clinical trials)

- Consider use of HL7 defined data types
- Consider the XML-ITS for using the data types
- Consider modeling using RIM objects for Clinical Trial space
- Consider adopting an NCI OID tree registered with HL7 for Cancer Center use
- Adopt CTS specification
- Take advantage of existing HL7 tooling